



OIPE

RAW SEQUENCE LISTING

DATE: 07/10/2002

PATENT APPLICATION: US/10/056,454A

TIME: 13:07:08

Input Set : A:\1627D.txt

Output Set: N:\CRF3\07102002\J056454A.raw

SEQUENCE LISTING

4 (1) GENERAL INFORMATION:

6 (i) APPLICANT: National Starch and Chemical Investment Holding Corporation

8 (ii) TITLE OF INVENTION: Improvements in or Relating to Plant Starch Composition

10 (iii) NUMBER OF SEQUENCES: 20

12 (iv) CORRESPONDENCE ADDRESS:

13 (A) ADDRESSEE: National Starch and Chemical Investment Holding Corporation

14 (B) STREET: 1000 Uniqema Blvd.

15 (C) CITY: Newcastle

16 (D) STATE: Delaware

17 (E) COUNTRY: United States of America

18 (F) ZIP: 19720

20 (v) COMPUTER READABLE FORM:

21 (A) MEDIUM TYPE: Floppy disk

22 (B) COMPUTER: IBM PC compatible

23 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

24 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

C--> 0 (vi) CURRENT APPLICATION DATA:

C--> 0 (A) APPLICATION NUMBER: US/10/056,454A

C--> 0 (B) FILING DATE: 25-Jun-2002

27 (2) INFORMATION FOR SEQ ID NO: 1:

29 (i) SEQUENCE CHARACTERISTICS:

30 (A) LENGTH: 57 base pairs

31 (B) TYPE: nucleic acid

32 (C) STRANDEDNESS: single

33 (D) TOPOLOGY: linear

39 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

41 AAGGATCCGT CGACATCGAT AATACGACTC ACTATAGGGA TTTTTTTTTT TTTTTT 57

43 (2) INFORMATION FOR SEQ ID NO: 2:

45 (i) SEQUENCE CHARACTERISTICS:

46 (A) LENGTH: 17 base pairs

47 (B) TYPE: nucleic acid

48 (C) STRANDEDNESS: single

49 (D) TOPOLOGY: linear

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

57 AAGGATCCGT CGACATC 17

59 (2) INFORMATION FOR SEQ ID NO: 3:

61 (i) SEQUENCE CHARACTERISTICS:

62 (A) LENGTH: 17 base pairs

63 (B) TYPE: nucleic acid

64 (C) STRANDEDNESS: single

65 (D) TOPOLOGY: linear

71 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/056,454A

DATE: 07/10/2002
TIME: 13:07:08

Input Set : A:\1627D.txt
Output Set: N:\CRF3\07102002\J056454A.raw

```
73 GACATCGATA ATACGAC 17
75 (2) INFORMATION FOR SEQ ID NO: 4:
77 (i) SEQUENCE CHARACTERISTICS:
78 (A) LENGTH: 20 base pairs
79 (B) TYPE: nucleic acid
80 (C) STRANDEDNESS: single
81 (D) TOPOLOGY: linear
87 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
89 CATCCAACCA CCATCTCGCA 20
91 (2) INFORMATION FOR SEQ ID NO: 5:
93 (i) SEQUENCE CHARACTERISTICS:
94 (A) LENGTH: 20 base pairs
95 (B) TYPE: nucleic acid
96 (C) STRANDEDNESS: single
97 (D) TOPOLOGY: linear
103 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
105 TTGAGAGAAG ATACCTAAGT 20
107 (2) INFORMATION FOR SEQ ID NO: 6:
109 (i) SEQUENCE CHARACTERISTICS:
110 (A) LENGTH: 20 base pairs
111 (B) TYPE: nucleic acid
112 (C) STRANDEDNESS: single
113 (D) TOPOLOGY: linear
119 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
121 ATGTTTCAGTC CATCTAAAGT 20
123 (2) INFORMATION FOR SEQ ID NO: 7:
125 (i) SEQUENCE CHARACTERISTICS:
126 (A) LENGTH: 20 base pairs
127 (B) TYPE: nucleic acid
128 (C) STRANDEDNESS: single
129 (D) TOPOLOGY: linear
135 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
137 AGAACAACAA TTCCTAGCTC 20
139 (2) INFORMATION FOR SEQ ID NO: 8:
141 (i) SEQUENCE CHARACTERISTICS:
142 (A) LENGTH: 20 base pairs
143 (B) TYPE: nucleic acid
144 (C) STRANDEDNESS: single
145 (D) TOPOLOGY: linear
151 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
153 GGGGCCCTGA ACTCAGCAAT 20
155 (2) INFORMATION FOR SEQ ID NO: 9:
157 (i) SEQUENCE CHARACTERISTICS:
158 (A) LENGTH: 20 base pairs
159 (B) TYPE: nucleic acid
160 (C) STRANDEDNESS: single
161 (D) TOPOLOGY: linear
167 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
169 CGTCCCAGCA TTCGACATAA 20
```

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/10/056,454A

DATE: 07/10/2002
 TIME: 13:07:08

Input Set : A:\1627D.txt
 Output Set: N:\CRF3\07102002\J056454A.raw

```

171 (2) INFORMATION FOR SEQ ID NO: 10:
173   (i) SEQUENCE CHARACTERISTICS:
174       (A) LENGTH: 26 base pairs
175       (B) TYPE: nucleic acid
176       (C) STRANDEDNESS: single
177       (D) TOPOLOGY: linear
183   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:
185 CTTGGATCCT TGAAGTACAG AATTG      26
187 (2) INFORMATION FOR SEQ ID NO: 11:
189   (i) SEQUENCE CHARACTERISTICS:
190       (A) LENGTH: 29 base pairs
191       (B) TYPE: nucleic acid
192       (C) STRANDEDNESS: single
193       (D) TOPOLOGY: linear
199   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:
201 TAACTCGAGC AACGCGATCA CAAGTTCGT      29
203 (2) INFORMATION FOR SEQ ID NO: 12:
205   (i) SEQUENCE CHARACTERISTICS:
206       (A) LENGTH: 3003 base pairs
207       (B) TYPE: nucleic acid
208       (C) STRANDEDNESS: single
209       (D) TOPOLOGY: linear
215   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:
217 GATGGGGCCT TGAAGTACAG AATTGACAC TCAGTTAGTT AACTGCCAT CACTTATCAG      60
219 ATCTCTATTT TTCTCTTAA TTCCAACCAA GGAATGAATA AAAAGATAGA TTTGTAAAAA      120
221 CCCTAAGGAG AGAAGAAGAA AGATGGTGTA TACACTCTCT GGAGTTCGTT TTCCTACTGT      180
223 TCCATCAGTG TACAAATCTA ATGGATTGAG CAGTAATGGT GATCGGAGGA ATGCTAATAT      240
225 TTCTGTATTC TTGAAAAAAC ACTCTCTTTC ACGGAAGATC TTGGCTGAAA AGTCTTCTTA      300
227 CAATTCGGAA TCCCGACCTT CTACAATTGC AGCATCGGGG AAAGTCCTTG TGCCTGGAAT      360
229 CCAGAGTGAT AGCTCCTCAT CCTCAACAGA TCAATTTGAG TTCGCTGAGA CATCTCCAGA      420
231 AAATTCCTCA GCATCAACTG ATGTAGATAG TTCAACAATG GAACACGCTA GCCAGATTAA      480
233 AACTGAGAAC GATGACGTTG AGCCGCTCAAG TGATCTTACA GGAAGTGTG AAGAGCTGGA      540
235 TTTTGCTTCA TCACTACAAC TACAAGAAGG TGATAAACTG GAGGAGTCTA AAACATTAAA      600
237 TACTTCTGAA GAGACAATTA TTGATGAATC TGATAGGATC AGAGAGAGGG GCATCCCTCC      660
239 ACCTGGACTT GGTGAGAAGA TTTATGAAAT AGACCCCTTT TTGACAACT ATCGTCAACA      720
241 CCTTGATTAC AGGTATTCAC AGTACAAGAA ACTGAGGGAG GCAATTGACA AGTATGAGGG      780
243 TGGTTTGGAA GCTTTTTCTC GTGGTTATGA AAGAATGGGT TTCACTCGTA GTGCTACAGG      840
245 TATCACTTAC CGTGAGTGGG CTCCTGGTGC CCAGTCAGCT GCCCTCATTG GGGATTTCAA      900
247 CAATTGGGAC GCAAATGCTG ACTTTATGAC TCGGAATGAA TTTGGTGTCT GAGAGATTTT      960
249 TCTGCCAAAT AATGTGGATG GTTCTCTGCG AATTCCTCAT GGGTCCAGAG TGAAGATACG      1020
251 TATGGACACT CCATCAGGTT TTAAGGATTC CATTCTGCT TGGATCAACT ACTCTTTACA      1080
253 GCTTCCTGAT GAAATTCAT ATAATGGAAT ATATATGAT CCACCCGAAG AGGAGAGGTA      1140
255 TATCTTCCAA CACCCACGGC CAAAGAAACC AAAGTCGGTG AGAATATATG AATCTCATAT      1200
257 TGAATGAGT AGTCCGGAGC CTAAAATTAA CTCATACGTG AATTTTAGAG ATGAAGTTCT      1260
259 TCCTCGCATA AAAAAAGCTT GGGTACAATG CGGTGCAAT TATGGCTATT CAAGAGCAT      1320
261 CTTATTATGC TAGTTTTGTT TATCATGTCA CAAATTTTTT TGCACCAAGC AGCCGTTTTG      1380
263 GAACGCCCGA CGACCTTAAG TCTTTGATTG ATAAAGCTCA TGAGCTAGGA ATTGTTGTTT      1440
265 TCATGGACAT TGTTCACAGC CATGCATCAA ATAATACTTT AGATGGACTG AACATGTTT      1500
267 ACGGCACAGA TAGTTGTTAC TTTCACTCTG GAGCTCGTGG TTATCATTGG ATGTGGGATT      1560

```

RAW SEQUENCE LISTING

DATE: 07/10/2002

PATENT APPLICATION: US/10/056,454A

TIME: 13:07:08

Input Set : A:\1627D.txt

Output Set: N:\CRF3\07102002\J056454A.raw

269	TCCGCTCTT	TAACATGGA	AACTGGGAGG	TACTTAGGTA	TCTTCTCTCA	AATGCGAGAT	1620
271	GGTGGTTGGA	TGAGTTCAAA	TTTGATGGAT	TTAGATTGGA	TGGTGIGACA	TCAATGATGT	1680
273	GTACTCACCA	CGGATTATCG	GTGGGATTCA	CTGGGAACATA	CGAGGAATAC	TTTGAGACTCG	1740
275	CAACTGATGT	GGATGCTGTT	GCTGATCTGA	TGCTGGTCAA	CGATCTTTAT	CATGGGCTTT	1800
277	TCCCAGATGC	AATTACCATT	GGTGAAGATG	TTAGCGGAAT	GCCGACATTT	TGTGTTCCCG	1860
279	TTCAAGATGG	GGGTGTTGGC	TTTGACTATC	GGCTGCATAT	GGCAATTGCT	GATAAATGGA	1920
281	TTGAGTTGCT	CAAGAAACGG	GATGAGGATT	GGAGAGTGGG	TGATATTGTT	CATACACTGA	1980
283	CAAAATAGAAG	ATGGTCGGAA	AAGTGTGTTT	CATACGCTGA	AAGTCATGAT	CAAGCTCTAG	2040
285	TCGGTGATAA	AACTATAGCA	TTCTGGCTGA	TGGACAAGGA	TATGTATGAT	TTTATGGGTC	2100
287	TGGATAGACC	GTCAACATCA	TTAATAGATC	GTGGGATAGC	ATTACACAAG	ATGATTAGGC	2160
289	TTGTAACAT	GGGATTAGGA	GGAGAAGGGT	ACCTAAATTT	CATGGGAAAT	GAATTCGGCC	2220
291	ACCCTGAGTG	GATTGATTTT	CCTAGGGCTG	AACAACACCT	CTCTGATGGC	TCAGTAATTC	2280
293	CCAGAAACCA	ATTCAGTTAT	GATAAATGCA	GACGGAGATT	TGACCTGGGA	GATCGAGAAT	2340
295	ATTTAAGATA	CCGTGGGTTG	CAAGAATTTG	ACCGGGCTAT	GCAGTATCTT	GAAGATAAAT	2400
297	ATGAGTTTAT	GACTTCAGAA	CACCAGTTCA	TATCACGAAA	GGATGGAAGG	GATAGGATGA	2460
299	TTGTAATTTGA	AAAAGGAAAC	CTAGTTTTTG	TCTTTAATTT	TCACTGGACA	AAAGGCTATT	2520
301	CAGACTATCG	CATAGCTGCG	CTGAAGCCTG	GAAAATACAA	GGTTGCCTTG	GACTCAGATG	2580
303	ATCCACTTTT	TGGTGGCTTC	GGGAGAATTG	ATCATAATGC	CGAATATTTT	ACCTTTGAAG	2640
305	GATGGTATGA	TGATCGTCCT	CGTTCAATTA	TGGTGTATGC	ACCTAGTAGA	ACAGCAGTGG	2700
307	TCTATGCACT	AGTAGACAAA	GAAGAAGAAG	AAGAAGAAGA	AGTAGCAGTA	GTAGAAGAAG	2760
309	TAGTAGTAGA	AGAAGAATGA	ACGAACCTGT	GATCGCGTTG	AAAGATTTGA	ACGCCACATA	2820
311	GAGCTTCTTG	ACGTATCTGG	CAATATTGCA	TTAGTCTTGG	CGGAATTTC	TGTGACAACA	2880
313	GGTTTGCAT	TCTTTCCACT	ATTAGTAGTG	CAACGATATA	CGCAGAGATG	AAGTGCTGAA	2940
315	CAAAAACATA	TGTAATATCG	ATGAATTTAT	GTCGAATGCT	GGGACGATCG	AATTCTTGCA	3000
317	GCC						3003
319	(2) INFORMATION FOR SEQ ID NO: 13:						
321	(i) SEQUENCE CHARACTERISTICS:						
322	(A) LENGTH: 2975 base pairs						
323	(B) TYPE: nucleic acid						
324	(C) STRANDEDNESS: single						
325	(D) TOPOLOGY: linear						
331	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:						
333	TTGATGGGCC	TTGAACCTCAG	CAATTTGACA	CTCAGTTAGT	TACACTCCTA	TCACTTATCA	60
335	GATCTCTATT	TTTTCTCTTA	ATTCCAACCA	GGGGAATGAA	TAAAAGGATA	GATTTTGTA	120
337	AACCCCTAAG	AGAGAAGAAG	AAAGATGGTG	TATATACTCT	CTGGAGTTGG	TTTTCTTACT	180
339	GTTCATCAG	TGTACAAATC	TAATGGATTG	AGCAGTAATG	GTGATCGGAG	GAATGCTAAT	240
341	GTTTCTGTAT	TCTTGAAAAA	GCACCTCTCT	TCACGGAAGA	TCTTGGCTGA	AAAGTCTTCT	300
343	TACAATTCGG	AATTCGGACC	TTCTACAGTT	GCAGCATCGG	GGAAGTCCCT	TGTGCTTGGA	360
345	ACCCAGAGTG	ATAGCTCCTC	ATCTCTCAAC	GACCAATTTG	AGTTCACTGA	GACATCTCCA	420
347	GAATTTCTCC	CAGCATCAAC	TGATGTAGAT	AGTTCAACAA	TGGAACACGC	TAGCCAGATT	480
349	AAAACCTGAGA	ACGATGACGT	TGAGCCGTC	AGTGATCTTA	CAGGAAGTGT	TGAAGAGCTG	540
351	GATTTTGCTT	CATCACTACA	ACTACAAGAA	GGTGGTAAAC	TGGAGGAGTC	TAAAACATTA	600
353	AATACCTCTG	AAGAGACAAT	TATTGATGAA	TCTGATAGGA	TCAGAGAGAG	GGGCATCCCT	660
355	CCACCTGGAC	TTGGTCAGAA	GATTGTGAA	ATAGACCCCT	TTTTGACAAA	CTATCGTCAA	720
357	CACCTTGATT	ACAGGTATTC	ACAGTACAAG	AACTGAGGGG	AGGCAATTGA	CAAGTATGAG	780
359	GGTGGTTTGG	AAGCTTTTCT	CGTGGTTATG	AAAAAATGGG	TTTCACTCGT	AGTGCTACAG	840
361	GTATCACTTA	CCGTGAGTGG	GCTCTGGTG	CCCAGTCAGC	TGCCCTCATT	GGAGATTTCA	900
363	ACAATTGGGA	CGCAAAATGCT	GACATTATGA	CTCGGAATGA	ATTTGGTGTC	TGGGAGATTT	960
365	TTCTGCCAAA	TAATGTGGAT	GGTTCTCTCG	CAATTCCTCA	TGGGTCAGGA	GTGAAGATAC	1020

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/056,454A

DATE: 07/10/2002

TIME: 13:07:08

Input Set : A:\1627D.txt

Output Set: N:\CRF3\07102002\J056454A.raw

```

367 GTATGGACAC TCCATCAGGT GTTAAGGATT CCATTCTCTGC TTGGATCAAC TACTCTTTAC 1080
369 AGCTTCTCTGA TGA AATTC A TATAATGGAA TATATTATGA TCCACCCGAA GAGGAGAGGT 1140
371 ATATCTTCCA ACACCCACGG CCAAAGAAAC CAAAGTCGCT GAGAATATAT GAATCTCATA 1200
373 TTGGATGAG TAGTCCGGAG CTTAAATTA ACTCATACGT GAATTTTGA GATGAAGTTC 1260
375 TTCCTCGCAT AAAAAAGCTT GGGTACAATG CGCTGCGAAT TATGGCTATT CAAGAGCATT 1320
377 CTTATTATGC TAGTTTGGT TATCATGTCA CAAATTTTTT TGCACCAAGC AGCCGTTTTG 1380
379 GAACGCCCGA CGACCTTAAG TCTTCGATTG ATAAAGCTCA TGAGCTAGGA ATTGTGTTC 1440
381 TCATGGACAT CGTTCACAGC CATGCAATCA ATAATACTTT AGATGGACTG AACATGTTTG 1500
383 ACGGCACCGA TAGTTGTTAC TTTCACCTCT GAGCTCGTGG TTATCATTGG ATGTGGGATT 1560
385 CCGCCTCTTT AACTATGGAA ACTGGGAGGT ACTTAGGTAT CTCTCTCAA ATGCGAGATG 1620
387 GTGGTTGGAT GAGTTCAAAT TTGATGGATT TAGATTGATG GGTGTGACAT CAATGATGTA 1680
389 TACTCACCA CCGATTATCG TGGGATTCAC TGGGAAC TAC GAGGAATACT TTGGACTCGC 1740
391 AACTGATGTG GATGCTGTTG TGTATCTGAT GCTGGTCAAC GATCTTATTC ATAGGCTTTT 1800
393 CCCAGATGCA ATTACCATTG GTGAAGATGT TAGCGGAATG CCGACATTTT GTATCCCGT 1860
395 TCAAGATGGG GGTGTTGGCT TTGACTATCG GCTGCATATG GCAATTGCTG ATAAATGGAT 1920
397 TGAGTTGCTC AAGAAACGGG ATGAGGATTG GAGAGTGGGT GATATTGTTC ATACACTGAC 1980
399 AAATAGAAGA TGGTCGGA AAGTGTGTTT ATACGCTGAA AGTCATGATC AAGCTCTAGT 2040
401 CGGTGATAAA ACTATAGCAT TCTGGCTGAT GGACAAGGAT ATGTATGATT TTATGGCTCT 2100
403 GGATAGACCG CCAACATCAT TAATAGATCG TGGGATAGCA TTGCACAAGA TGATTAGGCT 2160
405 TGTAACATATG GATTAGGAG GAGAAGGGTA CCTAAATTTT ATGGGAAATG AATTGCGCCA 2220
407 CCCTGAGTGG ATTGATTTC CTAGGGCTGA GCCACACCTT TCTGATGGCT CAGTAATTC 2280
409 CGGAACCAA TTCAGTTATG ATAAATGCAG ACGGAGATTT GACCTGGGAG ATGCAGAATA 2340
411 TTAAAGATAC CATGGGTTAC AAGAATTTGA CTGGGCTATG CAGTATCTTG AAGATAAATA 2400
413 TGAGTTTATG ACTTCAGAAC ACCAGTTTAT ATCAGGAAAG GATGAAGGAG ATAGGATGAT 2460
415 TGTATTTGAA AGAGGAAACC TAGTTTTCGT CTTTAATTTT CACTGGACAA ATAGCTATTC 2520
417 AGACTATCGC ATAGGCTGCC TGAAGCTGG AAAATACAAG GTTGTCTTGG ACTCAGAIGA 2580
419 TCCACTTTTT GGTGGCTTCG GGAGAATTGA TCATAATGCC GAATATTTCA CCTCTGAAGG 2640
421 ATCGTATGAT GATCGTCCCT GTTCAATTAT GGTGTATGCA CCTAGTAGAA CAGCAGTGGT 2700
423 CTATGCAC TAAGACAAAC TAGAAGTAGC AGTAGTAGAA GAACCCATTG AAGAATGAAC 2760
425 GAACTTGTGA TCGCGTTGAA AGATTGTAAC GTTACTTGGT CATCCACATA GAGCTTCTTG 2820
427 ACATCAGTCT TGGCGGAATT GCATGTGACA ACAAGGTTTG CAGTTCTTTC CACTATTAGT 2880
429 AGTCCACCGA TATACGAGA GATGAAGTGC TGAACAAACA TATGTAAAT CGATGAATTT 2940
431 ATGTCGAATG CTGGGACGAT CGAATTGCTG CAGCC 2975
433 (2) INFORMATION FOR SEQ ID NO: 14:
435 (i) SEQUENCE CHARACTERISTICS:
436 (A) LENGTH: 3033 base pairs
437 (B) TYPE: nucleic acid
438 (C) STRANDEDNESS: single
439 (D) TOPOLOGY: linear
443 (ix) FEATURE:
444 (A) NAME/KEY: CDS
445 (B) LOCATION: 145..2790
448 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:
450 TTGATGGGGC CTTGAACTCA GCAATTTGAC ACTCAGTTAG TTACTCTCCT ATCACTTATC 60
452 AGATCTCTAT TTTTCTCTT AATTCACAC AAGGAATGAA TAAAAGGATA GATTGTGAAA 120
454 AACCCCTAAGG AGAGAAGAAG AAAG ATG GTG TAT ACA CTC TCT GGA GTT CGT 171
455 Met Val Tyr Thr Leu Ser Gly Val Arg
456 1 5
458 TTT CCT ACT GTT CCA TCA GTG TAC AAA TCT AAT GGA TTC AGC AGT AAT 219

```

VERIFICATION SUMMARY

DATE: 07/10/2002

PATENT APPLICATION: US/10/056,454A

TIME: 13:07:09

Input Set : A:\1627D.txt

Output Set: N:\CRF3\07102002\J056454A.raw

L:0 M:249 C: Inserted Mandatory Field, [(vi) CURRENT APPLICATION DATA:]
L:0 M:249 C: Inserted Mandatory Field, [(A) APPLICATION NUMBER:]
L:0 M:249 C: Inserted Mandatory Field, [(B) FILING DATE:]